MODEL DESIGN IN THE PROSPECTOR CONSULTANT SYSTEM FOR MINERAL EXPLORATION*

Richard Duda, John Gaschnig and Peter Hart

ABSTRACT

Prospector is a computer consultant system intended to aid geologists in evaluating the favorability of an exploration site or region for occurrences of ore deposits of particular types. Knowledge about a particular type of ore deposit is encoded in a computational model representing observable geological features and the relative significance thereof. We describe the form of models in Prospector, focussing on inference networks of geological assertions and the Bayesian propagation formalism used to represent the judgmental reasoning process of the economic geologist who serves as model designer. Following the initial design of a model, simple performance evaluation techniques are used to assess the extent to which the performance of the model reflects faithfully the intent of the model designer. These results identify specific portions of the model that might benefit from "fine tuning", and establish priorities for such revisions. This description of the Prospector system and the model design process serves to illustrate the process of transferring human expertise about a subjective domain into a mechanical realization.

I. INTRODUCTION

In an increasingly complex and specialized world, human expertise about diverse subjects spanning scientific, economic, social, and political issues plays an increasingly important role in the functioning of all kinds of organizations. Although computers have become indispensable tools in many endeavors, we continue to rely heavily on the human expert's ability to identify and synthesize diverse factors, to form judgments, evaluate alternatives, and make decisions – in sum, to apply his or her years of experience to the problem at hand. This is especially valid with regard to domains that are not easily amenable to precise scientific formulations, i.e., to domains in which experience and subjective judgment plays a major role.

The precious resource of human expertise is also a fragile and transient one: the departure of a crucial expert from an organization may cause serious dislocations; senior people impart their knowledge to younger colleagues, but the demand for their talents may not leave sufficient time for such educational efforts.

* This work was supported in part by the Office of Resource Analysis of the U.S. Geological Survey under Contract No. 14-08-0001-15985, and in part by the Nonrenewable Resources Section of the National Science Foundation under Grant AER77-04499. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors, and do not necessarily reflect the views of either the U.S. Geological Survey or the National Science Foundation.

During recent years research in the field of artificial intelligence has produced effective new techniques for representing empirical judgmental knowledge and using this knowledge in performing plausible reasoning. The best known application of these techniques has been in the area of medical diagnosis, where computer programs have achieved high levels of performance (Pople et al., 1975; Shortliffe, 1976; Weiss et al., 1977; Yu, 1978; Szolovits, 1976). Other applications include planning experiments in molecular genetics (Martin 1977) and monitoring instruments in intensive care units (Fagan 1978). This paper concerns a similar computer program, called Prospector, that is being developed to help geologists in exploring for hard-rock mineral deposits. The characteristic of plausible reasoning shared by the domains of medical diagnosis and mineral exploration is common, to some degree, to many other diverse evaluation tasks as well. Hence the purpose of this paper is to illustrate, by a case study for the domain of mineral exploration, the general process of capturing and encoding human expertise into a mechanical realization.

II. OVERVIEW OF THE PROSPECTOR SYSTEM

The Prospector system is intended to emulate the reasoning process of an experienced exploration geologist in assessing a given prospect site or region for its likelihood of containing an ore deposit of the type represented by the model he or she designed. Here we use the term "model" to refer to a body of knowledge about a particular domain of expertise that is encoded into the system and on which the system can act. The empirical knowledge contained in Prospector consists of a number of such specially encoded models of certain classes of ore deposits. These models are intended to represent the most authoritative and up-to-date information available about each deposit class.

In Prospector's normal interactive consultation mode, the user is assumed to have obtained some promising field data and is assumed to desire assistance in evaluating the prospect. Thus, the user begins by providing the program with a list of rocks and minerals observed, and by inputting other observations expressed in simple English sentences. The program matches these data against its models, requests additional information of potential value for arriving at more definite conclusions, and provides a summary of the findings. The user can ask at any time for an elaboration of the intent of a question, or for the geological rationale for including a question in the model, or for an ongoing trace of the effects of his answers on Prospector's conclusions. The intent is to provide the user with many of the services that could be provided by giving him telephone access to a panel of senior economic geologists, each an authority on a particular class of ore deposits.

The performance of Prospector depends on the number of models it contains, the types of deposits modeled, and the quality and completeness of each model. Because the Prospector program is primarily a research project, its coverage is still incomplete. It currently contains five prospect-scale models, one regional-scale model, and one drilling site selection model. The prospect-scale models consist of a Kuroko-type massive sulfide model contributed by Charles F. Park, Jr., a Mississippi-Valley-type carbonate lead/zinc model contributed by Neil Campbell, a near-continental-margin porphyry copper model contributed by Marco T. Einaudi, a Komatiitic nickel sulfide model contributed by Anthony J. Naldrett, and a Western-states sandstone uranium model contributed by Ruffin I. Rackley. The regional scale model is a variation of Mr. Rackley's model. The drilling site selection model, for porphyry copper deposits, was contributed by Mr. Victor Hollister, and differs somewhat from the other models: it derives its inputs from digitized maps of geological characteristics, and produces as output a color-coded graphical display of the favorability of each cell on a grid corresponding to the input map. These models were selected for a variety of reasons, including their economic significance, the extent to which they are well understood scientifically, the availability of expert geologists who could collaborate with us in the model development, and the new research issues that their implementation would raise.

Each model is encoded as a separate data structure, independent of the Prospector system per se. Thus, the Prospector program should not be confused with its models. Rather, Prospector should be thought of as a general mechanism for delivering relevant expert information about ore deposits to a user who can supply it with data about a particular prospect or region. .

This paper describes briefly the process of developing and encoding such models for Prospector. General overviews of the technical principles are given in Hart, Duda and Einaudi (1978), mathematical aspects in Duda et al. (1976, 1978a), and detailed expositions in Duda et al. (1977, 1978b) and Hart, Duda and Konolige (1978).

III. FORMALISM FOR ENCODING EXPLORATION MODELS

A. Inference Networks of Assertions

For use in Prospector an ore deposit model must be encoded as a so-called inference network, a network of connections or relations between field evidence and important geological hypotheses. Since we sometimes do not wish to distinguish between evidence and hypotheses, we shall refer to either one as an assertion. To illustrate these ideas, we shall draw upon examples taken from M.T. Einaudi's porphyry copper model, which we shall denote by PCDA. Typical assertions in PCDA are "Hornblende has been pervasively altered to biotite" and "The alteration suggests the potassic zone of a porphyry copper deposit." The former would normally be thought of as field evidence, the latter as a geological hypothesis. A small portion of the PCDA inference network is shown in Figure 1. Here the terminal or "leaf" nodes correspond to field evidence asked of the user, while the other nodes represent hypotheses. The text in the boxes in Figure 1 is concise for reasons of graphical display; the actual questions asked of the user are more definitive.

Although assertions are statement that should be either true or false, in a given situation there is usually uncertainty as to whether they are true or false. Initially, the state of each assertion is simply unknown. As evidence is gathered, some assertions may be definitely established, whereas others may become only more

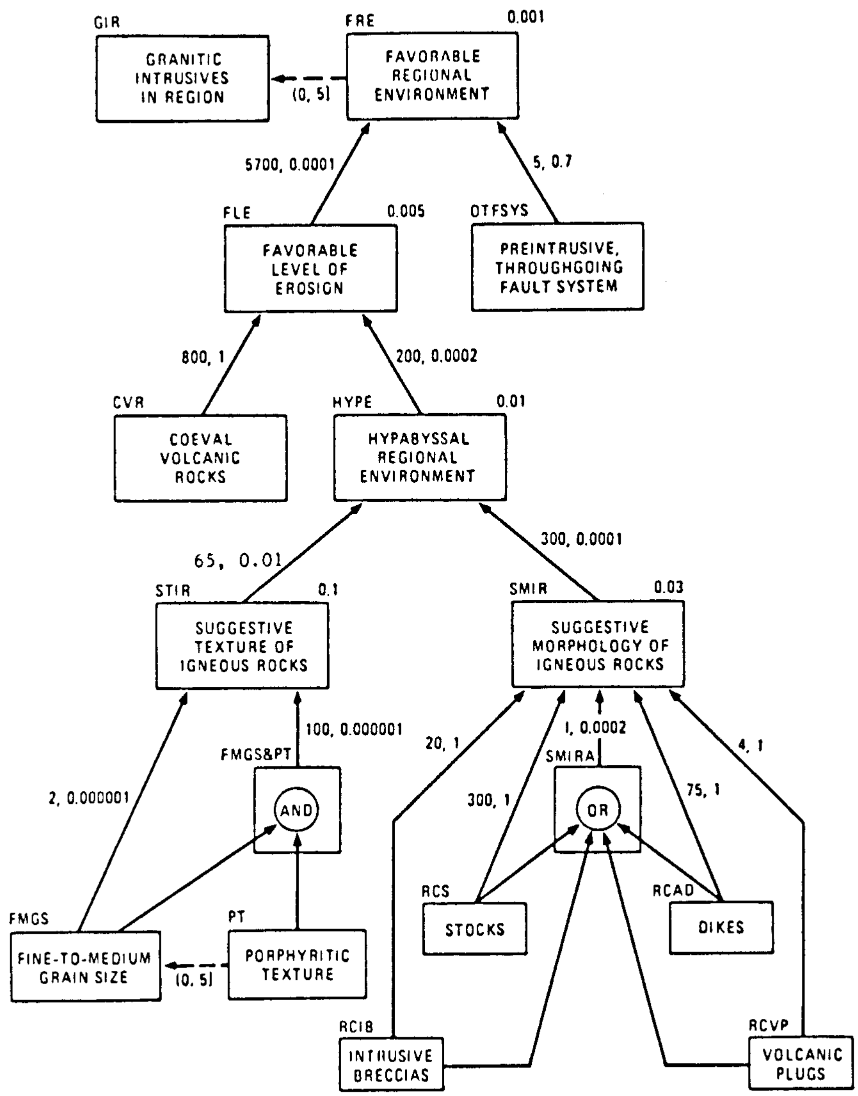

Figure 1. Portion of a Prospector model for porphyry copper deposits. See text for interpretation

or less likely. In general, we associate a probability value with every assertion. The "connections" in the inference network determine how a change in the probability of one assertion will affect those of other assertions.

The principal or top-level assertion in an inference network for a model is the assertion that the available evidence matches that particular model. To establish this assertion, it is usually necessary to establish several major factors. For example, to establish the top-level assertion in PCDA, we must establish the following hypotheses:

1. The petrotectonic setting is favorable for PCDA;
2. The regional evironment is favorable for PCDA;
3. There is an intrusive system that is favorable for PCDA.

Were any of these assertions field-observable evidence, they could be established merely by asking the user of the program whether they were true. However, since all of these factors are hypotheses, each must be further related to other factors. For example, the favorability of the petrotectonic setting can be established through the following three factors, each of which happens to be determinable (at least in principle) from observational evidence:

1. The prospect lies in a continental margin mobile belt;
2. The age of the belt is post-Paleozoic;
3. The prospect is subject to tectonic and magmatic activity related to subduction.

In general, the ore deposit models in Prospector have this type of hierarchical structure. The top-level assertion is determined by several major second-level assertions, each of which may be determined by third-level assertions, with this refinement continuing until assertions are reached that can be established directly from field evidence. This is illustrated in Figure 1, which shows graphically that portion of the PCDA model that describes the regional environment. In addition to this "top-to-bottom" development in terms of successive levels of assertions, the models also often exhibit a "left-to-right" organization in terms of spatial scale, from the petrotectonic setting on the left to the local details of mineralization and texture on the right. Exactly how these considerations interact is determined by the relations that exist among the assertions. The following section explains the nature of these relations and illustrates their occurrence in Figure 1.

B. Relations

Three basically different kinds of relations are used in Prospector to specify how a change in the probability of one assertion affects the probability of other assertions. We distinguish these as logical relations, plausible relations, and contextual relations.

1. *Logical Relations*. With logical relations, the truth (or falsity) of a hypothesis is completely determined by the truth (or falsity) of the assertions that define it. Such relations are composed out of the primitive operations of conjunction (AND), disjunction (OR), and negation (NOT). When several assertions must all be true for a hypothesis to be true, the hypothesis is the conjunction of the assertions. When the hypothesis is true if any of the assertions is true, the hypothesis is the disjunction of the assertions. Negation merely complements an assertion, interchanging truth and falsity. As an example of a logical relation, the PCDA model says that alteration of plagioclase is indicative of the barren-core zone if

1. Plagioclase has been altered to
   a. albite
   or
   b. minor sericite (or both)

and

2. Plagioclase has not been altered to major epidote.

Other examples of logical relations are shown in Figure 1.

Of course, in general we do not know whether the assertions are true, but can only estimate a probability or degree of belief that they are true. With logical relations, to compute the probability of a hypothesis from the probability of its component assertions we employ the fuzzy-set formulas of Zadeh (1965). Using these formulas, the probability of a hypothesis that is defined as the logical conjunction (AND) of several pieces of evidence equals the minimum of the probability values corresponding to the evidence. Similarly, a hypothesis defined as the logical disjunction (OR) of its evidence spaces is assigned a probability value equal to the maximum of those values assigned to the evidence spaces. One property of this procedure is that it often gives no "partial credit." In particular, if all but one of the assertions have been established, but the user can not even guess about the last, then the probability of their conjunction often remains at the value it had when the states of none of the assertions were known. This may be the appropriate conclusion. When it is not, one has the option of using plausible relations.

2. *Plausible Relations.* With plausible relations, each assertion contributes "votes" for or against the truth of the hypothesis. This would be expressed by relating the assertions to the hypothesis through a set of plausible inference rules. Each rule has an associated rule strength that measures the degree to which a change in the probability of the evidence assertion changes the probability of the hypothesis. This change can be positive or negative, since as assertion can be either favorable or unfavorable for a hypothesis. As with all parts of a model, these rule strengths are obtained by interviewing an authority on the corresponding class of ore deposits. Initially he may express the strengths in verbal terms, such as "rather discouraging" or "very encouraging." This is ultimately translated into numerical terms (as shown in Figure 1), the changes in probability being computed in accordance with the rules of Bayesian probability theory, as outlined below and described in detail in Duda et al. (1977).

Prospector's plausible reasoning scheme is based on Bayesian decision theory (Raiffa, 1968), exploiting an elementary theorem of probability known as Bayes' rule. For our purposes, the so-called "odds-likelihood" form of the rule is most convenient. This form relates three quantities involving an evidence assertion E and a hypothesis assertion H: the prior odds O(H) on the hypothesis, the posterior odds O(H | E) on the hypothesis, given that E is observed to be present, and a measure of sufficiency LS. Then Bayes' rule can be stated as

$$O(H \mid E) = LS * O(H) \tag{1}$$

Odds and probabilities are freely interchangeable through the simple relation $O = P / (1 - P)$, where P denotes probability, and hence $P = O / (1 + O)$. The sufficiency measure LS is a standard quantity in statistics called the likelihood ratio, and is defined by

$$LS = \frac{P(E|H)}{P(E|\sim H)} \tag{2}$$

where ~H means "not H."

Equation (1) prescribes a means for updating the probability (or odds) on H, given that the evidence E is observed to be present. An inference rule for which LS is large means that the observation of E is encouraging for H – in the extreme case of LS approaching infinity, E is sufficient to establish H in a strict logical sense. On the other hand, if LS is much less than unity, then the observation of E is discouraging for H, inasmuch as the observation of E diminishes the odds on H.

A complementary set of equations describes the case in which E is known to be absent, that is, when ~E is true. In this case, we can use Bayes' rule to write

$$O(H \mid \sim E) = LN * O(H) \tag{3}$$

where

$$LN = \frac{P(\sim E \mid H)}{P(\sim E \mid \sim H)} \tag{4}$$

The quantity LN is called the necessity measure. If LN is much less than unity, the known absence of E transforms neutral prior odds on H into very small posterior odds in favor of H. In the extreme case of LN approaching zero, E is logically necessary for H. On the other hand, if LN is large, then the absence of E is encouraging for H.

Hence to define an inference rule

IF E
THEN (to degree LS, LN) H,

the model designer must articulate E and H, and must supply numerical values for LS, LN, and O(H).

In general, the user may not be able to state that E is either definitely present or definitely absent. In this case, the updating formulas (1) and (3) cannot be applied directly, but can be extended to accommodate the uncertainty in the evidence. The extension used in Prospector involves a linear interpolation between the extremes of E's being definitely present or definitely absent. See Duda (1976, 1977) for details. The user expresses his certainty about E on an arbitrary –5 to 5 scale, where 5 denotes that the evidence is definitely present, –5 denotes that it is definitely absent, 0 indicates no information, and intermediate values denote degrees of certainty.

We illustrate this plausible inference scheme with examples taken from Figure 1. The two numbers associated with each inference rule in Figure 1 are its LS and LN values, respectively. The number appearing above each box representing a nonterminal node is the prior probability of that assertion's being true. For example, the figure indicates that the existence of stocks is a more highly sufficient indicator of "suggestive morphology of igneous rocks" (i.e., LS = 300) than is the existence of either dikes, intrusive breccias, or volcanic plugs (i.e., LS = 75, 20, and 4, respectively). Similarly, "favorable level of erosion" (FLE) is a highly sufficient and highly necessary factor for establishing "favorable regional environment" (i.e., LS = 5700 and LN = 0.0001), whereas the existence of a "preintrusive throughgoing fault system" (OTFSYS) is only mildly sufficient and mildly necessary for establishing "favorable regional environment." Hence the positive (LS) or negative (LN) votes of FLE are weighted much more heavily than those of OTFSYS.

The section of the model concerned with establishing "suggestive morphology of igneous rocks" (SMIR) illustrates how logical and plausible relations can be combined as building blocks to accomplish the intent of the economic geologist designing the model. This section of the PCDA model can be described as follows. "There are four positive indicators for establishing a suggestive morphology for igneous rocks (SMIR), namely intrusive breccias, stocks, dikes, and volcanic plugs. Each of these factors contributes independently to establishing SMIR, although to differing degrees. The absence of any one of these four factors individually is unimportant [i.e., LN = 1 for those rules]. However, if it is known that none of these factors is present [implying that the disjunction node SMIRA is false], then the probability of a suggestive morphology of igneous rocks is essentially zero [LN = 0.0002 for SMIRA]." In defining an inference network for a model, the object is to induce the model designer to articulate such statements, and then to translate the statements into network constructions.

To see how the effect of a piece of evidence propagates upward through the model, suppose that the user has indicated only that intrusive breccias are present, but this is definite. This fact multiplies the odds of SMIR by a factor of 20, hence raising its probability from 0.03 to 0.382. (The prior odds on SMIR are 0.03 / (1 – 0.03) = 0.030927, giving posterior odds on SMIR equal to 20 * 0.030927 = .61855, which corresponds to a probability of 0.61855 / (1+0.61855) = 0.382.) This in turn increase the odds on HYPE by a factor of 300 weighted by the degree to which SMIR has increased from its prior probability, i.e, by the factor 300 * (0.382–0.03) / (1–0.03) = 108.866. Hence the posterior probability of HYPE is 0.52373, which in turn increases the odds of FLE by a factor of 200 * (0.52373–0.01) / (1–0.01) = 103.78, giving a posterior probability for FLE of 0.34276. The propagation continues in this manner upward through the network.

It should be noted that Prospector expresses its conclusions to the user on the same –5 to 5 certainty scale that the user employs to express his certainty about evidence requested by the system. Prospector maps internal probability values to external certainty scores in a piecewise linear fashion, such that the posterior certainty is proportional to the difference between the posterior probability and the prior probability. For example, since the prior probability of

FLE is 0.005, a posterior probability of 0.34276 corresponds to a posterior certainty of 5 * (0.34276−0.005) / (1−0.005) = 1.697. Similarly, a posterior probability of 0.001 corresponds in this case to a posterior certainty of 5 * (0.001 − 0.005) / 0.005 = −4. See Shortliffe (1975) for a description of the subjective certainty scale used in the MYCIN medical diagnosis system.

3. *Contextual Relations.* It sometimes happens that assertions cannot be considered in an arbitrary order, but must be considered in a particular sequence. For example, one should determine that there is a relevant continental margin mobile belt before considering its age. This is more than a matter of preference, since it would be meaningless for the program to ask about the age of a non-existent belt.

To treat such situations we employ the third class of relations, contextual relations. In general, we use contexts to express a condition that must be established before an assertion can be used in the reasoning process. In the above example, the existence of a continental margin mobile belt would be specified as a context for asking about the age of the belt. Thus, before inquiring about the age, the system would employ all its resources to establish the existence of the belt, and would not ask about its age unless the probability of the belt were greater than its initial value.

Contextual relations are also used when one assertion is geologically significant only if another assertion has already been established. In such instances it would not be nonsensical to ask the former question without first establishing the latter, but it is the case that the former evidence is geologically irrelevant without the latter to establishing a match to the model. Two such instances are depicted by dashed arrows in Figure 1. In one of these instances, the entire "favorable regional environment" section of PCDA model will not be pursued unless it has first been determined that there are granitic intrusives in the region.

IV. OVERVIEW OF THE MODEL DEVELOPMENT PROCESS

Although the development and encoding of a model for Prospector is not a routine process, it does progress through several distinct phases whose general nature can be described. The four most important phases are summarized below.

A. Initial Preparation

Model development is a cooperative enterprise involving an exploration geologist, who is an authority on the type of deposit being modeled, and a computer scientist who understands the operation of the Prospector system. The first step in developing a model is one of introducing the exploration geologist (model designer) to the inference network formalism, and introducing the computer scientist (model implementor) to the general nature of the class of deposits being modeled. In particular, this includes the identification of several known deposits that should fit the model well, and several known deposits that may fit partially, but that lack certain important characteristics. These specific cases help to establish the various factors that must be taken into account.

B. Initial Design

The initial design of the inference network is the most creative phase of the process. It requires the identification of the various assertions, the organization of the assertions into a hierarchical structure (as illustrated in Figure 1), the determination of the types of relations (logical, plausible, and contextual) that exist among assertions, and the estimation of values for the parameters (the voting strengths and initial probabilities). The magnitude of this task depends upon the size and complexity of the model being developed; as a point of reference, the smallest model currently in Prospector contains 28 assertions and 20 inference rules, while the largest contains 212 assertions and 133 inference rules. The initial design is usually facilitated by considering factors in the "top-down" and "left-to-right" sequence described earlier. Delicate refinement is best avoided at this time, since subsequent revision often causes significant sections of the model to be reorganized, enlarged, or otherwise modified.

In addition to the connections between assertions exhibited directly by the inference network, there are connections that exist because of the geological meaning of the assertions. For example, the statement that there are sulfide minerals is obviously related to the statement that there is pyrite in quartz veins; assertion of the latter implies the former, and denial of the former denies the latter. Recognition of such connections within a model avoids redundant or foolish questioning; recognition of such connections between different models allows the program to consider more than one deposit class at a time. Prospector can automatically recognize many of these assertions if each assertion is properly articulated. This articulation, which is described in more detail in Duda (1978b), should also be completed during the initial design.

C. Installation and Debugging of the Model

At the end of Phase B, the model exists in a "pencil-and-paper" form. To be incorporated into the program, the encoding must be given a formal description. This is done through the use of a model description language (see Duda, 1978b). The details of this language are not particularly important here. However, the task itself is important; upon its completion the program can be run, and accidental blunders or bugs can be corrected. In addition, the program can produce a questionnaire for the model that is useful in gathering data for subsequent testing and revision.

D. Performance Evaluation and Model Revision

Given the questionnaire data for a number of actual deposits, it is possible to make a serious quantitative evaluation of how well particular deposits match the model. In our experience, this evaluation inevitably exposes various shortcomings of the model as encoded, requiring revision of the work done in Phases B and C. Some care must be exercised here to avoid "overfitting" the model to the data. In general, the goal is to produce a model that can discriminate different types of deposits without losing the ability to generalize, so as to allow for the variations one would expect in new situations. Achievement of that goal currently remains as much an art as a science. The following section describes in some detail the use of simple performance evaluation techniques as an aid to refining a model.

V. USE OF PERFORMANCE EVALUATION IN REFINING A MODEL

To demonstrate that the performance of an expert knowledge-based system is (or is not) comparable to that of the experts it emulates, it is useful to subject the system to an appropriate objective evaluation. The simple performance evaluation experiments reported in this section serve several purposes: (1) to provide an objective, detailed, quantitative measure of the current performance of a model; (2) to pinpoint those sections of the model that are not performing exactly as intended, thereby establishing priorities for future revisions; (3) to assess consistency of performance across different exploration sites.

We now evaluate a model for a class of porphyry copper deposits (PCDA) designed by Prof. Marco Einaudi of Stanford University. Input data were available for three test cases, namely, the known deposits called Yerington (Nevada), Bingham (Utah), and Kalamazoo (Arizona), each of which is considered an exemplar of the PCDA model.[2] On the –5 to 5 certainty scale described earlier, the overall certainty scores computed by Prospector are 4.769 for the Yerington deposit, 4.721 for Bingham, and 4.756 for Kalamazoo, indicating a good match of these sites to the PCDA model.

To show performance in detail, we give below the hierarchical structure of the major sections of the PCDA model. Included at the right in this enumeration is the total number of questions that may be asked by Prospector for each of the major sections of the model, thus showing the relative distribution of these questions. (The questions in the FAMR section may be asked several times during a consultation session, once for each geographically distinct zone within the prospect area. Each such zone has relatively homogeneous geological characteristics, as determined by the user.)

	Total Number of Questions Defined in PCDA Model (Version 2)
Porphyry Copper deposit, type A (PCDA)	81
Favorable petrotectonic setting (FPTS)	4
Favorable regional environment (FRE)	9
Favorable PCDA intrusive system (FPCDAIS)	68
Favorable composition in differentiated sequence (FCDS)	4
Favorable intrusive system (FIS)	9
Favorable alteration and mineralization relations (FAMR)	56

[2]The questionnaire input data used in the present tests are reported in Duda et al., 1978b, pp.185-93.

As a calibration exercise, Prof. Einaudi offered a target value for the certainty score that should be assigned to each of the three deposits for each of the major components of the model listed above, based on the values in the input data set for each prospect site. The target values are given either in the form of a single number (on a –5 to 5 scale), or as two numbers establishing an upper and lower bound on a certainty interval. The estimates are listed in Table 1 on the left for each site in turn, with the scores as determined by execution of Prospector recorded on the right. (We informed Prof. Einaudi of the values on the right only after he had given us those on the left.)

Yerington Deposit

Name of Model Node	Einaudi's Estimate	Prospector Score
PCDA	4.5 to 5.0	4.769
FPTS	4.5 to 5.0	4.528
FRE	4.5	4.540
FPCDAIS	4.5 to 5.0	4.787
FCDS	5	4.524
FIS	5	4.744
FAMR	4.5 to 5.0	4.225

Bingham Deposit

Name of Model Node	Einaudi's Estimate	Prospector Score
PCDA	4.5	4.721
FPTS	3.5 to 4.0	4.449
FRE	4.0 to 4.5	4.829
FPCDAIS	4.5 to 5.0	4.729
FCDS	5	2.407
FIS	5	4.744
FAMR	4.0	4.225

Kalamazoo Deposit

Name of Model Node	Einaudi's Estimate	Prospector Score
PCDA	4.0 to 4.5	4.756
FPTS	4.0 to 4.5	4.449
FRE	3.5	1.784
FPCDAIS	4.5 to 5.0	4.791
FCDS	5	4.722
FIS	5	4.744
FAMR	4.0	4.225

Table 1. Prospector Scores for Several Levels of the PCDA Model (Version 2)

The data in Table 1 show that Prospector scores each of these sections of the model with high certainty for each site, with the exception that model node FCDS for the Bingham deposit and model node FRE for the Kalamazoo deposit are scored somewhat lower. In most cases shown in Table 1 Prospector agrees very closely with Prof. Einaudi's estimate. These conclusions can be expressed quantitatively by first identifying the values in Table 1 with a concise notation, then defining a simple formula for the relative error of Prospector in predicting Prof. Einaudi's estimates. Thus:

Let C(X, Y, Z) = Certainty score given to model node Z by agent X for site Y,
where X denotes either Prospector or Einaudi

For example, C(Prospector, Yerington, FPCDAIS) = 4.787. When Einaudi gave an interval of certainty values instead of a single value, we use the mid-point of the interval as the value of C. Then an error measure is given by

$$E(Y, Z) = \frac{C(\text{Einaudi}, Y, Z) - C(\text{Prospector}, Y, Z)}{C(\text{Einaudi}, Y, Z)}$$

For example, E(Yerington, FPCDAIS) = (4.75 – 4.787) / 4.75 = –0.008, meaning that Prospector's prediction is accurate to within 0.8% in this case. Since Table 1 gives values for seven nodes of the model for each of three known deposits, we can compute the value of E for 21 different instances. For 5 of the 21 data points Prospector predicted Einaudi's estimate to within 1%, while 15 of the 21 data points show agreement to within 10%. The grand average over the 21 data points is 10.3%. For convenience, we list these 21 values of E in Table 2, expressed as percentages.

	Yerington	Bingham	Kalamazoo	Average of Absolute Values
PCDA	–.3 %	–4.9 %	–11.9 %	5.7 %
FPTS	4.7	–18.6	–4.7	9.3
FRE	–.9	–13.6	49.0	21.2
FPCDAIS	–.8	.4	–.9	.7
FCDS	9.5	51.9	5.6	22.3
FIS	5.1	5.1	5.1	5.1
FAMR	11.1	5.6	5.6	7.6
Average of Absolute values:	4.1	14.3	11.8	10.3

Table 2. Relative Error (E) of Prospector Scores as Predictors of Einaudi's Estimates (derived from data in Table 1)

Inspection of Table 2 indicates that efforts to revise the PCDA model should focus on the FRE and FCDS sections. When such revisions are completed, an updated version of Table 2 will indicate the extent to which the revisions achieved the objectives that motivated them.

The small number of cases tested, and the fact that all the present test cases are exemplars of the PCDA model, and the fact that the model designer himself supplied the input data concerning the test cases, are limitations; the present tests are more necessary than sufficient conditions for good performance. Despite these limitations, the preliminary results reported here have proved useful in the ongoing model refinement process. More extensive performance evaluation results are reported in Duda et al. (1978b).

VI. REMARKS

This paper has outlined the typical procedures used to develop an exploration model for the Prospector system. We have described the inference network and Bayesian propagation scheme underlying Prospector models, and we have illustrated the use of simple performance evaluation techniques in "fine-tuning" a model systematically. Our experience indicates that the model design process inherently requires feedback. Although different problem solving domains differ in many details, we believe the process of constructing Prospector-like plausible reasoning systems follows certain general patterns and stages of development such as are described here. Hence we have presented a concrete case study, in the domain of mineral exploration, that may credibly suggest what might be expected in attempts to apply a similar methodology to other domains of expertise.

Besides the running program, there appear to be several other benefits to this type of expert system approach. The model design process challenges the model designer to articulate, organize, and quantify his expertise. Without exception, the economic geologists who have designed Prospector models have reported that the experience aided and sharpened their own thinking on the subject matter of the model. In addition, most of the geologists we know who have had experience with Prospector have remarked about its potential value as an educational tool. In this regard, the models in the system contain explicit, detailed information synthesized from the literature and the experience of expert explorationists, together with explanatory text that can be obtained upon request. Furthermore, a typical consultation session with Prospector costs only about $10 at current commercial computer rates.

VII. ACKNOWLEDGEMENTS

Prospector would not exist without the talents and efforts contributed by many economic geologists and computer scientists. Among these, we are particularly indebted to the economic geologists Alan N. Campbell, Neil Campbell, Marco T. Einaudi, Victor F. Hollister, Anthony J. Naldrett, Charles F. Park, Jr., and Ruffin I.Rackley, and to the computer scientists Phyllis Barrett, Kurt Konolige, Nils J.Nilsson, Rene Reboh, Jonathan Slocum, and B.Michael Wilber.

REFERENCES

Duda, R.O., P.E. Hart and N.J. Nilsson (1976), 'Subjective Bayesian methods for rule-based inference systems,' *Proc. National Computer Conference*, (AFIPS Conference Proceedings, Vol. 45), pp. 1075-1082, 1976.

Duda, R.O., P.E. Hart, N.J. Nilsson, R. Reboh, J. Slocum and G.L. Sutherland (1977), 'Development of a computer-based consultant for mineral exploration,' Annual Report, SRI Projects 5821 and 6415, SRI International, Menlo Park, California, October 1977.

Duda, R.O., P.E. Hart, N.J. Nilsson, and G.L. Sutherland (1978a), 'Semantic network representations in rule-based inference systems,' in *Pattern Directed Inference Systems*, D.A. Waterman and F. Hayes-Roth (Eds.), pp. 203-221, Academic Press, New York, 1978.

Duda, R.O., P.E.Hart, P.Barrett, J.Gaschnig, K.Konolige, R.Reboh and J.Slocum (1978b), "Development of the Prospector system for mineral exploration," Final Report, SRI Projects 5821 and 6415, SRI International, Menlo Park, California, October 1978.

Fagan, L.M. (1978), "Ventilator Manager: A program to provide on-line consultative advice in the intensive care unit." Heuristic Programming Project Memo HPP-78-16, Dept. of Computer Science, Stanford University, September 1978.

Hart, P.E., R.O. Duda, and K. Konolige (1978), "A computer-based consultant for mineral exploration," Second Semiannual Report, SRI Project 6415, SRI International, Menlo Park, California, April 1978.

Hart, P.E., R.O. Duda, and M.T. Einaudi (1978), "A computer-based consultation system for mineral exploration," to appear in *Computer Methods for the 80's*, Society of Mining Engineers of the AIME, 1978.

Martin, N., P. Friedland, J. King, and M. Stefik (1977), "Knowledge Base Management for Experiment Planning in Molecular Genetics," *Proc. of 5th International Joint Conference on Artificial Intelligence*, pp. 882-887, Cambridge, Mass., August 1977.

Pople, H.E., Jr., J.D. Myers, and R.A. Miller (1975), "DIALOG: A model of diagnostic logic for internal medicine," *Proc. of 4th International Joint Conference on Artificial Intelligence*. pp. 848-855, Tbilisi, Georgia, USSR, 1975.

Raiffa, H. (1968), *Decision Analysis*, Addison-Wesley Publishing Co., New York, 1968.

Shortliffe, E.H. (1976), *Computer-Based Medical Consultations: MYCIN*, American Elsevier Publishing Co., New York, 1976.

Shortliffe, E.H., and B.G. Buchanan (1975), "A model of inexact reasoning in medicine," *Math. Biosci.*, Vol. 23, pp. 351-379, 1975.

Szolovits, P., and S.G. Pauker (1976), "Research on a medical consultation system for taking the present illness," *Proc. of 3rd Illinois Conference on Medical Information Systems*, pp. 299-320, University of Illinois at Chicago Circle, November 1976.

Weiss, S.M., C.A. Kulikowski, and A. Safir (1977), "A model-based consultation system for the long-term management of glaucoma," *Proc. of 5th International Joint Conference on Artificial Intelligence*, pp. 826-833, MIT, Cambridge, Massachusetts, 1977.

Yu, V.L., et al. (1978), "Evaluating the performance of a computer-based consultant," Heuristic Programming Project Memo HPP-78-17, Dept. of Computer Science, Stanford University, September 1978.

Zadeh, L.A. (1965), "Fuzzy sets." *Information and Control*, Vol. 8, pp. 338-353, 1965.